Technical Report 1025

Virtual Reality Psychophysics: Forward and Lateral Distance, Height, and Speed Perceptions With a Wide-Angle Helmet Display

Robert H. Wright
U.S. Army Research Institute

April 1995

ARMY RESEARCH INSTITUTE

United States Army Research Institute
for the Behavioral and Social Sciences

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE

3. REPORT TYPE AND DATES COVERED
Final Jul 92 - Apr 94

4. TITLE AND SUBTITLE
Virtual Reality Psychophysics: Forward and Lateral Distance, Height, and Speed Perceptions With a Wide-Angle Helmet Display

5. FUNDING NUMBERS
62785A
791
2211
H01

6. AUTHOR(S)
Wright, Robert H.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army Research Institute for the Behavioral and Social Sciences
ATTN: PERI-IR
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

8. PERFORMING ORGANIZATION REPORT NUMBER
ARI Technical Report 1025

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Institute for the Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
--

11. SUPPLEMENTARY NOTES
--

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release;
distribution is unlimited.

12b. DISTRIBUTION CODE
--

13. ABSTRACT *(Maximum 200 words)*

Psychophysics of a color, high resolution, very wide angle, virtual reality type of helmet-mounted display were investigated. Subjects used a joystick to set their viewpoint within a computer-generated image database to requested target values in forward and lateral distance, height, and speed. Test factors for each type of perception included helicopter flying experience, replications, 3-D with familiar objects or 2-D texture visual databases, relative or absolute perceptions, viewpoint motion rail row and column offsets, increasing or decreasing change in target values, and six target values. Median forward distance and speed perceptions were 41% of simulated physical stimuli, 50% for lateral distance, and 72% for height. These accuracies contrast with typical real-world accuracies for similar ranges of about 90% for distance, height, and speed. Main effort differences between most of the test factor levels were highly significant for all four types of perceptions.

14. SUBJECT TERMS
Psychophysics, Virtual reality, Perceptual accuracy, Distance, Height, Speed, Computer imagery

15. NUMBER OF PAGES
42

16. PRICE CODE
--

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

U.S. ARMY RESEARCH INSTITUTE
FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Director

Technical review by

John A. Dohme
John E. Stewart
Dennis C. Wightman
Bob G. Witmer

Accesion For		
NTIS CRA&I	☑	
DTIC TAB	☐	
Unannounced	☐	
Justification		
By		
Distribution /		
Availability Codes		
Dist	Avail and / or Special	
A-1		

NOTICES

DISTRIBUTION: Primary distribution of this report has been made by ARI. Please address correspondence concerning distribution of reports to: U.S. Army Research Institute for the Behavioral and Social Sciences, ATTN: PERI-POX, 5001 Eisenhower Ave., Alexandria, Virginia 22333-5600.

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

Technical Report 1025

Virtual Reality Psychophysics: Forward and Lateral Distance, Height, and Speed Perceptions With a Wide-Angle Helmet Display

Robert H. Wright
U.S. Army Research Institute

Rotary-Wing Aviation Research Unit
Charles A. Gainer, Chief

Training Systems Research Division
Jack H. Hiller, Director

U.S. Army Research Institute for the Behavioral and Social Sciences
5001 Eisenhower Avenue, Alexandria, Virginia 22333-5600

Office, Deputy Chief of Staff for Personnel
Department of the Army

April 1995

Army Project Number
20262785A791

Education and Training Technology

FOREWORD

The U.S. Army plans to increase use of computer-generated imagery for simulation in initial and combat readiness training. Computer-generated imagery is the key technology in virtual reality, which is being explored for potential uses in training and mission rehearsals. How well visual perceptions with computer-generated imagery compare with real-world perceptions will be a major factor determining simulation and virtual reality effectiveness. The accuracy of visual perceptions with computer-generated imagery, however, remains almost totally undefined after nearly three decades of application in simulation. Difficulties in conducting research with simulators designed for training is a major reason for the failure to define their perceptual accuracies in comparison to the real world.

The Simulator Training Research Advanced Testbed for Aviation (STRATA) at the Rotary-Wing Aviation Research Unit (RWARU) of the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) has a very high quality virtual reality type of helmet display and is designed for research. STRATA was configured and used to obtain thousands of psychophysical observations per subject on distance, height, and speed perceptions. This allowed high statistical confidence on the effects of the factors tested.

The results indicate that distance, height, and speed perceptions with STRATA's very high quality virtual reality type of computer-generated imagery are substantially less than simulated physical values or real-world perceptions. The results suggest that substantial perceptual underestimations may be expected with existing and at least near-term computer-image generation technology. These perceptual underestimations may limit attainment of the highest levels of combat proficiency if training is limited to only that conducted in simulators which use computer-image generators. The results also suggest existing and developmental night vision systems that use similar digitized image formats might produce similar perceptual underestimations.

This psychophysical research was initiated internally by the ARIRWARU as the first effort of a series to objectively quantify the simulation fidelity of STRATA. It will be applied by the ARIRWARU in planning future STRATA research, and as a knowledge resource for application to issues concerning aviation simulator training and development.

EDGAR M. JOHNSON
Director

ACKNOWLEDGMENTS

The author gratefully acknowledges the programming support and assistance of Rande Doty in data collection, reduction, and statistical analysis, and of Larry Murdock in statistical analysis. My thanks to John Stewart for running some of the subjects.

CAE, Inc., site personnel headed by Nick Donker are thanked for their capable support in reconfiguring the STRATA simulator for this atypical psychophysical testing requirement. Rolf Beutler programmed the test conditions and sequences, and Fred Zalzal the test control monitor displays and data recording. Yves Provencher poured the subject helmet liners and aligned the helmet displays.

Majors Bill Barker and Dale Weiler are thanked for coordinating the scheduling of pilots to serve as subjects. A special thanks is extended to the subject pilots for their dedicated motivation in trying to provide accurate perceptual estimates over the many hours of testing.

VIRTUAL REALITY PSYCHOPHYSICS: FORWARD AND LATERAL DISTANCE, HEIGHT, AND SPEED PERCEPTIONS WITH A WIDE-ANGLE HELMET DISPLAY

EXECUTIVE SUMMARY

Requirement:

How well visual perceptions with computer-generated visual systems compare with real-world perceptions remains almost totally undefined. Such comparisons will determine the effectiveness and acceptance of computer-imaging systems in many applications for virtual reality, simulator training, and helicopter systems simulator-based research and development (R&D). Yet psychophysical research on visual perceptions with computer-generated visual systems is quite limited. Differences in visual perceptions between computer-synthesized visual worlds and the real world are significant factors that should be considered in planning and conducting simulator training and research, but have not been available. There is an unmet need, therefore, for psychophysical data on various types of visual perceptions with computer-generated imagery for simulator training research, system development R&D, and operational training and virtual reality applications.

Research objectives were to quantify psychophysically accuracy of forward and lateral distance, height, and speed perceptions with the pilot station virtual reality type of helmet display used in ARI's STRATA training research simulator. Factors the literature suggests might influence accuracy of these perceptions were included in the research design to assess their impact. Accuracy of these perceptions was to be related to comparable real-world perceptual data available in the literature.

Procedure:

Perceptual estimates were made by subjects viewing a computer-generated terrain image in a head-slaved, color, high-resolution, very wide angle helmet-mounted virtual reality-type display. Subjects used a joystick to adjust their viewing point to a sequence of six target values along computer-defined invisible rails through space. These rails ran parallel with the direction of the type of perception being estimated. Nine rails offset from the reference cue in a 3 by 3 matrix were used for distances and height perceptions, and four vertically offset rails used for speed. One visual condition consisted only of ground texture, roads, and five skyscraper-type towers, one of which was used as the reference cue. The second visual condition

added to the first trees and other familiar 3-D objects, such as trucks on the road. Absolute and, for distances and height, relative perceptions were made for both increasing and decreasing sequences of target values along each rail. Tower depth (43 m) was used as the reference cue for relative distance and height perceptions. This length also was depicted as the height and base of two adjoining white triangles painted on the lower right corner of the towers. Pilots both with and without helicopter flying experience were used as subjects. Errors found for perceptions with the computer-generated image were compared with similar perceptions in the real world.

Findings:

Perceptions for forward distance are 41% of the median simulated physical stimuli, 50% for lateral distance, 72% for height, and 41% for speed perceptions. These substantial perceptual underestimates contrast with typical real-world perceptions of about 90% for all four types of perceptions. Psychophysical power function exponents were 1.00 for forward distance, 0.91 for lateral distance, 0.88 for height, and 1.03 for speed. An exponent of one indicates a change in perception proportional to that in the physical stimulus; an exponent less than one a decreasing ratio of change in perception with increasing magnitude of the physical stimulus. There were highly significant main effect differences between most of the test factor levels for all four types of perceptions, and numerous significant interactions.

Perceptions were more accurate (6 to 13%) for pilots with helicopter flying experience than for those with none. Relative perceptions were considerably more accurate (14 to 24%) than absolute perceptions. Perceptions were more accurate (7 to 15%) when target values were increasing than when they were decreasing. Forward (4%) and lateral (3%) distance perceptions were slightly more accurate for the 3-D than for the texture visual data-base, but no significant visual database differences were found for height and speed perceptions. Perceptual accuracy decreased (10 to 14%) as height of viewpoint rails increased, and also as rail lateral offset from the reference cue increased (3 to 9%).

The magnitude production psychophysical methodology used appears to be a sensitive approach for determining the impact of visual system design and task characteristics on perceptual performance. Psychophysical evaluations of other simulator visual systems and other flying and fighting-related tasks appears to be warranted. Such evaluations would provide confirmation of these results and contribute to definition and insight into actual perceptual capabilities and limitations of simulator/virtual reality visual systems, which heretofore have been lacking.

Utilization of Findings:

The findings suggest that substantial perceptual underestimation of distance, speed, and height should be expected with computer-generated simulator or virtual reality visual systems. These perceptual errors could promote flying the simulator too fast and landing short and high. Simulator instructors and students should be made aware of these tendencies for perceptual errors with simulator visual systems. Similar error tendencies may be expected with virtual reality visual systems. The general psychophysical methodology used in this experiment should be considered for defining the perceptual characteristics of other computer-generated visual imaging systems.

VIRTUAL REALITY PSYCHOPHYSICS: FORWARD AND LATERAL DISTANCE, HEIGHT, AND SPEED PERCEPTIONS WITH A WIDE-ANGLE HELMET DISPLAY

CONTENTS

LIST OF TABLES

LIST OF FIGURES

VIRTUAL REALITY PSYCHOPHYSICS: FORWARD AND LATERAL DISTANCE, HEIGHT, AND SPEED PERCEPTIONS WITH A WIDE-ANGLE HELMET DISPLAY

Introduction

Psychophysical research on visual perceptions with computer generated visual systems is quite limited. How well visual perceptions of computer generated visual imagery compare with real-world perceptions remains almost totally undefined. Yet, such comparisons largely will determine the effectiveness and acceptance of such imaging systems in many applications for virtual reality, training, and simulator-based advanced system development.

This research psychophysically examines some of the primary self-motion spatial perceptions involved in helicopter flight control. It investigates forward and lateral distance, height, and speed perceptions with a high resolution, computer generated, color virtual reality type of helmet display.

Simulator Psychophysical Research

Simulator psychophysical research has two primary areas of emphasis: display hardware and scene content. Psychophysical criteria and research issues for visual display hardware were reviewed and defined by Kraft, Anderson & Elworth (1980). Kraft & Anderson (1980) investigated two of the issues, on display joints and scene inserts. For target-to-background luminance ratios above 10, increasing resolution of a targeting display from .95 to .5 arc min per picture element increased aircraft plan view aspect recognition ranges by 28%, rather than the 90% suggested by the resolution ratios (Kennedy, Berbaum, Collyer, May & Dunlap, 1988). Increasing resolution from 5 to 1.5 arc min was found to increase vehicular detection and identification ranges by 30 to 60%, rather than the 333% factor suggested by the resolution ratios (Barrette et al., 1990). Target detection ranges and combat flying tasks both improved as display luminance increased from 2 to 51 cd/m^2. Combat flying task performance improved as lateral field of view increased from 87 to 127 degrees, and binocular viewing increased head-on aircraft detection ranges by 18% over monocular viewing (Barrette et al., 1990). Palmer & Petitt (1977) found angular size estimates of 200% of actual for a distant triangle in a computer generated collimated night scene, and 150% for an uncollimated night scene. Gilinsky (1955) found 300% estimates for the same task outdoors in daylight.

More scene content research indicates improved flight performance for vertical objects than for objects coplanar with the ground (Buckland, Edwards & Stephens, 1981; Martin &

Rinalducci, 1983; Lintern, Thomley-Yates, Nelson, & Roscoe, 1987). The highest detail or least spacings used in ground texture and between vertical objects provides better performance (Kleiss & Hubbard, 1993; Buckland, Monroe & Mehrer, 1980; Buckland et al., 1981). However, the realism of object appearance has not been found to affect performance (Kleiss & Hubbard, 1993).

Realistic appearing 1.5, 4.6 and 10.7 m high objects (pine and oak trees) had no effect on altitude change perceptions in comparison to inverted geometric tetrahedrons (Kleiss & Hubbard, 1993). Increasing object density from 1 to 13/km^2 improved both percent correct and reaction time for altitude change perceptions, and increasing density to 51/km^2 resulted in further improvement in reaction time. Texture on the terrain surface also improved altitude change perceptions (Kleiss & Hubbard, 1993).

The best small jet landing performance was found for the shortest runway grid spacing investigated of only 1.2 m by Buckland, Monroe & Mehrer (1980). The least checker spacing of 67 m and the presence of vertical objects produced better performance in jets trying to fly at a 15 m height over simulated rolling terrain (Buckland, Edwards & Stephens, 1981). Holding 30 m or 61 m altitudes and avoiding terrain crashes was better with 3-D than with 2-D visual cues, with 457 m than with 1372 m cue spacing, and with higher contrast cues (Martin & Rinalducci, 1983). However, for a higher speed, McCormick, Smith, Lewandowski, & Preskar (1983) found 46 m ground texture squares to result in slightly higher altitudes than 91 m squares, and a mixture of 3-D objects slightly lower altitudes than single types of objects, or none. A landscape scene that contained buildings, roads and rectangular fields gave better performance than a schematic grid pattern for both training and transfer of a dive bombing task (Lintern et al., 1987).

Since its introduction, Stevens (1957) power law exponent (the "n" in the formula $Y = kX^n$) has been a primary metric used for psychophysical scaling of the relationship between perceptual magnitude (Y) and physical stimulus magnitude (X). The logarithmic form of the formula usually is used: $\log Y = \log k + n(\log X)$. The k in the formula is a scaling constant which is usually reported to reflect units of measurement. When the logarithm of Y is plotted against the logarithm of X, a 45° straight line results when the ratio of perceptual to physical magnitude is a constant. This 45° straight line has a slope of 1.0, and graphically reflects a power function exponent, n, of 1.0. When the ratio of perceptual to physical magnitude decreases as physical magnitude increases, n and the slope of the log-log plot will be less than one. When the ratio increases, n and the slope will be greater than one.

An exponent of 1.0 is a necessary, but not sufficient, condition for reflecting accurate perceptions. If perceptions are a constant 5% of stimulus magnitudes, for example, an exponent and slope of 1.0 will be found even though the perceptions are highly inaccurate. If such a constant perceptual error exists, it will shift perceptual magnitudes by that constant, in the same manner that the constant k shifts plotting levels to reflect the units of measurement. The combination of these two constant shifts will be reflected as a vertical shift in the slope line of the log-log plot. Misinformation, that an exponent of 1.0 represents accurate perceptions, has somehow permeated throughout a lot of the psychological literature of authors who are not specialized in psychophysical research (for example, Kling & Riggs, 1971).

De Maio & Brooks (1982) and De Maio, Rinalducci, Brooks & Brunderman (1983) investigated the effects of vertical object density and detail on the accuracy of altitude perceptions, using the Stevens exponent n as their psychophysical measure. They recorded simulator visual imagery as photographs, as dynamic video tape motion segments, and as static video tape views. The imagery was recorded at eight altitudes in 50 foot increments from 50 to 400 feet above ground level. The photographic and video tape records of various visual databases at the eight altitudes were viewed by groups of pilots on a large classroom projection screen. For each image or dynamic image segment, each pilot estimated the height above ground.

The power function exponents obtained from these estimations varied from a low of 0.20 to a high of 0.84 for the various visual databases. The higher exponents were obtained for the visual databases with the highest object density and detail. De Maio et al. (1983) concluded that the knee in the exponent versus object density curve at 0.7 (90% of the exponent asymptote just above 0.8), corresponding to an object density of about 5 objects per km^2, represented the point of diminishing returns in object density for height perceptions. For the highest exponent of 0.82 that De Maio & Brooks found for the recorded imagery, however, they report mean estimated altitudes of 25 ft for 50 ft (50%) actual simulator imagery height, and 257 ft for 400 ft (64%) actual. These end point heights result in a power function exponent of 1.12 rather than 0.82. Middle heights or individual variability must have had substantial effect, therefore, for their 0.82 result to occur. Although not reported for other heights, their exponent would have required estimated altitudes of around 75% of actual at 100 and 150 ft heights, estimates of less than 50% at 300 and 350 ft, or some similar combination of higher percentage estimates at lower heights, and lower estimates at the higher heights.

An in-simulator flying validation experiment by De Maio et al. (1983) found mean height perceptions to be about 80% of actual for three higher density and detail visual databases which had exponents near 0.8 in the above research. For a visual database with a 0.5 exponent mean, height perception accuracy of 60% of actual was found, and for the 0.2 exponent database, perceptions averaged 50% of actual heights. However, for two of the three visual databases with exponents near 0.8, pilots received 15-minute practice sessions with full instruments. This practice could be expected to enhance perceptual calibration for all three of these similar visual databases. The more accurate height perceptions with them may largely be due to this perceptual calibration. As conducted, the validation experiment results do reflect a positive linear relationship between psychophysical power function exponent and perceptual accuracy in height perception.

The method used and conclusions reached by De Maio & Brooks (1982) and De Maio et al. (1983) are based on erroneous assumptions about the meaning of the power function exponent with respect to perceptual accuracy. De Maio & Brooks, citing the experimental psychology textbook of Kling & Riggs (1971), state "A power (slope) of 1.0 indicates perfectly accurate estimate of altitude." This is not true, as indicated above. A power (slope) of 1.0 only reflects no change in the ratio between perceptual and physical stimulus magnitudes. A power (slope) of 1.0 will indicate perfectly accurate estimation only if perceptual and physical magnitudes are equal. This seldom will be the case for perceptions. Therefore, the conclusions concerning the power function exponent made by De Maio & Brooks and De Maio et al., are relatively meaningless with respect to perceptual accuracy. The power functions they determine for altitude estimates with the various visual databases, however, are valid results.

Upon initial casual reading, the results and conclusions of De Maio & Brooks (1982) and De Maio et al. (1983) appeared to be the most relevant in the literature to this proposed psychophysical research on simulator distance, height and speed perceptions. Upon the more detailed review described above, however, only the few observations they report on actual accuracy or percent accuracy, are considered to be germane results.

Real-World Psychophysical Research

Real world psychophysical research data on the type of perceptions used in this research were reviewed to derive single estimates of perceptual accuracy and power function exponents. These estimates were desired to allow comparisons with the four types of perceptions over the ranges used for them in this research. Where not reported and the data allowed, they were used to calculate power function exponents.

Typical distance perceptions were found to be about 90% of actual distances, and the power function exponent to be 1.00 (Gibson & Bergman, 1954; Gilinsky, 1951; Teghtsoonian & Teghtsoonian, 1970; Galanter & Galanter, 1973; Burney, 1977; Denz, Palmer & Ellis, 1980; Fine & Kobrick, 1981). Primary weight was given to the pretesting data of Gibson & Bergman for 92 subjects, which were far more than used by any other researchers. Calculations from their data indicated average perceptions to be 91.3% of actual distances, and the power function exponent was calculated to be exactly 1.00.

Real world research on helicopter height perceptions (Ungs & Sangal, 1990; Armstrong, Hofmann, Sanders, Stone & Bowen, 1975) indicate average perceptions are 87% of actual heights, and the power function exponents calculated from their data are 1.00. Ungs & Sangal found mean perceptions over land of 102% of actual heights if ascending to target heights of 7.6 to 61 m, but 76% of actual if descending to the same target heights. Their results suggest ascending or descending to target values is a factor which should be controlled in psychophysical research test design.

Real world research on speed perception by Salvatore (1968) and Armstrong et al. (1975) agree in indicating typical perceptions for normal vision should be about 90% of actual speeds, and power function calculations with their data indicate exponents of 1.00.

Definitive psychophysical data on simulator or virtual reality visual system distance, height and speed perceptions were not found. The few research results available tend to be preliminary non-definitive findings for jet aircraft that are not directly relevant to combat helicopter heights and speeds, or the visual databases used for helicopter simulation. There is a need, therefore, to quantify the accuracy of basic perceptions which result from viewing computer generated imagery, and the effects of factors which might influence that accuracy. The objective of this research is to respond to that need. Using the very high quality computer generated imagery of the STRATA, accuracy of perceptions of forward distance, lateral distance, height, and speed will be quantified psychophysically. Effects of a selected set of factors on these perceptions also will be assessed.

Method

A magnitude production psychophysical procedure without feedback was used in which subjects adjusted their helmet display viewpoint with joysticks to requested relative or absolute target values in forward and lateral distance, height and speed. Viewpoint motions were constrained to move in only the axis of perception along computer defined invisible rails (straight lines

through space). Nine of these rails were defined in a geometrically spaced matrix of viewpoint offsets in x, y or z from the visual reference, for each axis of distance or height perception. Four rails were defined in a vertical column beside a road for speed perceptions.

Subjects

Six male subjects were used. Three were Army helicopter pilots with 550, 1000 and 2300 hours of flying experience. The other three had only light fixed wing flying experience (170, 280 and 340 hours). The latter three had been selected for and were about to begin their Army helicopter training. All subjects had normal or corrected-to-normal visual acuity or better.

Apparatus

Subjects viewed a computer generated (Evans and Sutherland ESIG 1000) visual scene through a color helmet display from an AH-64 attack helicopter rear pilot station used as a simulator. It was enclosed by black drapes which reduced the surrounding dimmed laboratory lighting to scotopic levels. The usual computer synthesized replacement of the canopy framing in the helmet display image was turned off. But masking of the terrain view by blanking fuselage and glareshield areas to black level was retained. The helmet display allowed normal head movement, with slight attenuation from weight reduction lines and fiber-optic ropes used for transmitting images to the helmet. Helmet display instantaneous field of view was 65° vertical by 125° horizontal, with a high resolution center inset of 18 by 25°. Horizontal overlap of 38° existed at the vertical center of the images for the two eyes, decreasing in 41° circular arcs to 6.5° at the top and bottom. Non-stereo binocular viewing of partially identical (in the overlap area) infinity focus images was used.

Background images had 5, and the inset 1.5 arc min TV line resolution in both axes, for a 10% contrast modulation input. Display maximum white luminance capability exceeded 103 cd/m^2, and contrast 50:1. For the colors, textures and shading of features used in the visual database, however, typical contrast and luminance were only about 10 to 30% of system capabilities. Inset eye tracking was not functional for this testing, so the inset was fixed at the center of the display, and therefore was always viewed with full overlap. A blending algorithm was used at the edges of the inset, which usually resulted in no awareness of these edges for normal viewing. The helmet image continuously tracked the current angles and translations of the helmet. The simulated helicopter fuselage was fixed facing north for all distance and height perceptions, and fixed facing south for speed.

Target values to be estimated were presented on a CRT display in the upper center of the instrument panel, which was clearly visible through the blanked to black level area of the helmet display. For relative perceptual estimates, target values were presented as a single decimal fraction or integer multiple. For absolute distance or height estimates, target values were presented with units labels on the top line in feet, on the middle line in yards, and on the lower line in meters. For speed estimates units labeled target values were provided in statute miles per hour on the top line, and on the lower line in knots. The different units were intended to allow subjects to use those with which they were most familiar. All normal flight instruments and helmet display symbology were turned off.

Identical spring centered joysticks (Measurement Systems Model 546) were mounted on the right and left sides of the seat outside the armorplate. The right joystick controlled viewpoint fore-aft position and speed by fore-aft inputs, and lateral positions by right-left inputs. The left joystick was mounted with a 60° forward tilt. Up-down viewpoint positions were controlled by similar inputs with it, and right-left inputs used to control yaw during pre-test training free-flight. The trigger switch of the control used for viewpoint control was pressed to initiate a trial, and a pushbutton on top of the opposite control used to enter estimates for target values during a trial. Joystick responses were defined so non-helicopter pilots could rapidly and precisely adjust their position or speed. The center +/- 2% was a non-response deadband. Stick voltages beyond this band (minus the deadband) were squared and then multiplied by a gain which resulted in maximum viewpoint rates of 329 km/hr for distance and height estimation, and maximum acceleration response of 4.6 m/s^2 or 0.47 g for speed estimation. An integration time constant of one second was used for all control responses except switch inputs.

<u>Visual Databases</u>

Both visual databases consisted of a visually unlimited expanse of flat ground surface textured in shades of green. The detail area (see Figure 1) of the textured visual database added several roads and five tall skyscraper-type towers with their east walls aligned on X = 0. Shadows were not used for any type of object. The east-west visual database axis is labeled X, the north-south axis Y, and the height axis Z. All database coordinates and object dimensions are given in meters. The database X-Y-Z 0-0-0 and perceptual reference coordinate was at the southeast ground level corner of the south tower. A 9 m wide road with center at X = -194 extended infinitely southward from Y = 22. It turned 45° clockwise at Y = 22 to run northeastward between the two south towers. Another road ran east-west just north of the towers at Y = 989.

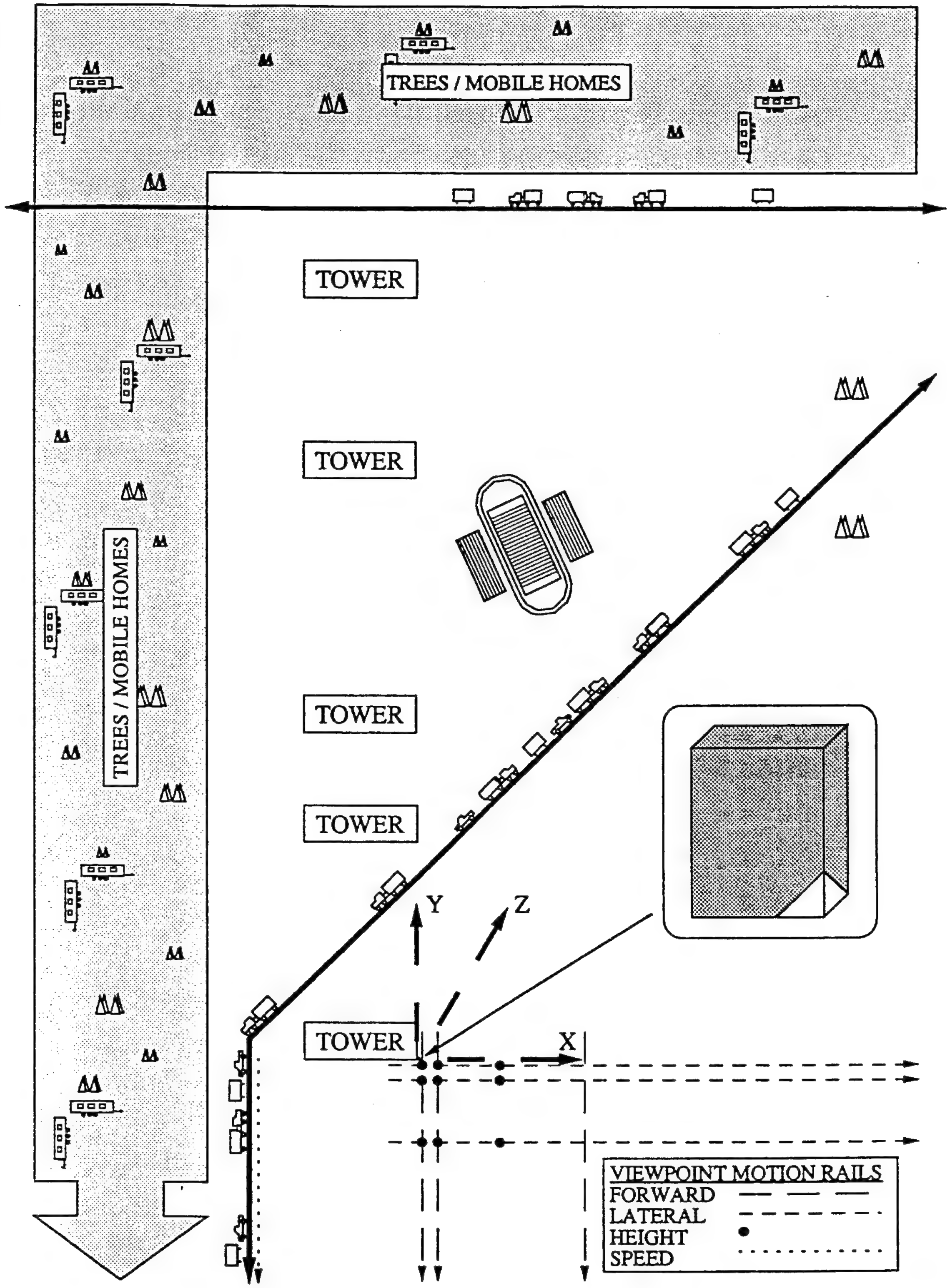

Figure 1. Detail area for psychophysical estimates. Ground surfaces within and beyond gaming area were depicted in a green colored texture pattern. Three forward and lateral rails (height 6, 24, and 98 m), and four speed rails (6, 12, 24 and 48 m) exist for each rail depicted in this plan view. Height rails rise from the page.

The towers were colored concrete gray, were 129 m wide, and 43 m deep along the Y coordinate. From south to north their heights were 701, 823, 729, 945 and 834 m, with respective spacings between them 215, 86, 258, and 172 m. At the southeast corner of each tower adjacent white 45° right triangles were drawn on the east and south walls with their 90° angle at the corner at ground level (see Figure 1 inset). Their high point was at the corner 43 m above the ground, and their bases extended 43 m along the ground. Relative perception references were towers depth, and triangle heights and bases, all being 43 m. The large towers and the white triangles were used to ensure that distinct cues for perceptual reference, and distinct relative perception cues, were available at the longest expected viewing distances and from all viewing angles.

The other (3-D with familiar objects) visual database added to the texture database several types of 3-D objects which are depicted schematically in Figure 1. The objects included variable size four-tree "groves," trucks on the roads and billboards beside them, mobile homes, and a high school type of football stadium.

The trees were not simple geometric objects. They were of an evergreen type, had a texture pattern which gave them a somewhat realistic appearance, and allowed viewing the background between the "branches." Tree groves consisted of four trees planted in a square pattern, with the grove edge-to-edge dimensions equal to tree height. Tree width and the spacing between them in a grove were one-third of tree height. Tree groves were scaled proportionally in sixteen equally spaced sizes between 5 and 37 m for their maximum dimensions. Two maximum size tree groves with north-south and east-west aligned edges were located just north and south of the 45° road with centers at X = 514, Y = 851, and X = 514, Y = 630.

Tree groves and mobile homes were placed within a continuous 200 m wide area (X = -237 to -437) on the west side of the detailed database. This area of tree groves and mobile homes extended on indefinitely to the south with the adjacent road (center at X = -194). A tree grove of randomly determined size and edge alignment was randomly located along each 75 m north-south interval of the area, and randomly along its width. The center of a mobile home (18L x 4W x 3.3H) with length oriented east-west was located 30 m south of the center of every fourth tree grove. A second adjacent mobile home with length oriented north-south was located with its center 14 m south and 14 m west of the center of the first mobile home. The same tree grove and mobile home area extended between Y = 1032 to 1232 along the north edge of the detail area from X = -437 to +613, but with mobile homes placed south of the adjacent tree grove.

Three truck types were used. One was an 18-wheel semi type with box trailer, the second the same with a tank trailer (both 18L x 2.4W x 4H overall), and the third a full size pickup truck (5.5L x 2W x 1.8H). Billboards 6 m wide by 4 m high were placed only on the west or north sides of roads, where they always faced the subject. For contrast with the ground texture, a range of light bright colors, white and silver, were used for the trucks, and their wheels were black. The billboard frames were colored light gray. Truck or billboard objects were randomly placed along each 75 m interval of the roads, with a restriction of no overlapping at interval ends of adjacent trucks in the same lane. The type of object in each interval, and its heading and color if a truck, were randomly determined. The stadium included a U.S. football field yard-marked outline with goal posts, a track around the field, and spectator stands on each side. The sidelines ran at an azimuth of 155°. The center of the stadium was located at X = 132, Y = 628.

No roads or 3-D objects existed in the area where viewpoint motion was allowed to the south of the south tower and east of the north-south road (see Figure 1). Ground texture was the only visual cue inside this area. Actual viewpoint motions extended well below the figure for forward distance perception trials, and well to the right of it for lateral distance trials. Ground texture was the only visual cue on the east (left looking south) side of the road during all speed estimates, and also on the west side during perceptions with the textured visual database.

Experimental Design

Experimental design in terms of data collection consisted of the factors outlined in Table 1, with the within factors listed in order of increasing cycling rate. Each type of perception was analyzed separately, however, rather than using perception as an additional factor. Factor levels are indicated in parentheses, and listed in the order tested for within factors (except for rows and columns as indicated below). Counterbalancing was not used. It should be noted this results in a full half-replication with the 3-D with familiar objects visual database prior to any perceptual estimates being required with the texture only visual database. The three-letter abbreviations are used for factor reference in some of the results later in the report.

Table 1

Experimental Design Factors and Factor Levels

Between							
EXP	Helicopter Experience (None; Experience)						
Pilots	(N1; N2; N3; E1; E2; E3;)						
Within							
REP	Replication (1; 2)						
VIS	Visual Database (3-D with Familiar Objects; Texture)						
BAS	Basis of Estimate (Relative; Absolute)						
Type	of Perception (Forward; Lateral; Height; Speed)						
ROW	Row of Viewpoint Offset (6; 24; 98 m)						
COL	Column of Viewpoint Offset (6; 24; 98 m [Forward 195])						
SGN	Sign of Target Value Change (Increasing; Decreasing)						
VAL	Target Values	(#1;	#2;	#3;	#4;	#5;	#6)
	Relative ratio:	(.25	.5	1	2	4	8)
	Relative (m):	(11	22	43	86	172	344)
	Absolute (m):	(8	15	30	61	152	305)
	Speed (km/hr):	(24	48	72	97	121	145)

Data were collected as a full factorial design, except that (a) 195 rather than the 98 m column offset was used for viewpoint rails for forward distance perceptions, (b) for speed only absolute perceptions were used, and (c) four row height offsets tested in the order 49, 24, 12 and 6 m, and a single column offset of 10 m east of road center, were used for speed. The double offset of the third column for forward distance perceptions, was used to minimize in one column the perspective shrinking of the primary relative reference cue (towers' east wall depth), to a single picture element or less at the longer viewing distances. For data analyses, an a priori decision was made to analyze the four types of perceptions separately. This was based on the logically different nature of these perceptions, the different factor levels for speed, and limited computer memory.

Table 2 indicates the fixed rail testing sequence for distance and height perceptions. This sequence resulted in a progression from least to most rail vector offset from the 0-0-0 point of visual reference, except for the center rail. Rows varied in Z and columns in X for forward distance perceptions, in Z and Y for lateral distance perceptions, and in X and Y for height perceptions.

Table 2

Sequence of Testing of Viewpoint Motion Rails

	Column		
	1	2	3
Row			
3	4	7	9
2	2	5	8
1	1	3	6

Prior to the test week, subjects had a helmet liner poured and helmet display attachment brackets boresighted. A pair of subjects were tested for 10 to 12 hours each in alternating sessions over a period of five days. Testing schedules were adjusted around other simulator uses. Most sessions were about 50 minutes in duration, but varied from 20 to 80 minutes. The first test morning began with a desktop briefing on the perceptual tasks and testing procedures, followed by about one hour of training on them in the cockpit and helmet display. Training ended with a free-flight period in which subjects were prompted to visually inspect at very close range, from different viewing angles, all 3-D objects in the visual database. Subjects were informed in the briefing that the total length of the 18-wheelers was 60 feet. No other dimensional information or feedback on accuracy of magnitude productions was provided to subjects during training or testing.

Image unblanking at start of trial. Forward and lateral distance, and height perceptual estimates were conducted with the simulated fuselage heading fixed to north. Image unblanking occurred for forward and lateral distance perceptions at the start of a trial, at an eyepoint on each rail, which was three times the initial target value from the perceptual reference corner of the south tower, but opposite in sign. In visual database coordinates, unblanking occurred for relative forward distance perceptions at Y = +32, and at Y = +23 for absolute perceptions. For lateral distance perceptions, unblanking was at X = -32 for relative, and at X = -23 for absolute perceptions. The image unblanked for all height perceptions one meter above the ground. For speed perceptions, the image unblanked with the fuselage facing south at Y = 0 and X = -184, 10 m east of the center of the north-south road. No limits to eyepoint motion existed in either direction along a rail. Subjects could go below ground level during height perceptions or free-flight, and several did so a few times during training or in early testing.

Procedures for distance and height perceptions. The subject began the first trial for forward distance estimation by pressing the trigger switch of the right joystick used for X-Y viewpoint motion control. This unblanked the image at current helmet angles. As the image unblanked the first of the sequence of increasing target values to be estimated (#1, equal to .25 relative or 8 m absolute) appeared on the panel CRT. The subject used fore-aft joystick input to move the eyepoint rearward along the rail to a location estimated to match the target value in distance south of the perceptual reference (the south wall of the south tower). When this location was attained, the subject entered the estimate by pressing the push button on top of the opposite (left) joystick. The next target value (#2) then appeared on the CRT, and the subject moved the eyepoint rearward to and entered the estimate for it. This process continued through estimation of the sixth and farthest target value. Upon entering the estimate for it, the image blanked to a uniform gray for 3 seconds as the eyepoint location was reset to start the sequence of decreasing target values to be estimated on the trial.

The reset location was at a distance on the rail three times the largest #6 target value, or at Y = -1032 m for relative perceptions and Y = -914 m for absolute perceptions. As the image unblanked, the first decreasing sequence target value to be estimated (#6) appeared on the CRT, which was the same as the last value of the increasing sequence. The subject used the joystick to move the eyepoint forward (or rearward) to a location estimated to match the target value in distance south of the reference, and entered the estimate. The next target value (#5) then appeared on the CRT, and an estimate was made for it. Estimation then continued through the last decreasing sequence target value (#1). Upon entering this last estimate for the rail, the image was blanked to gray and the eyepoint location reset without unblanking to the starting location for the next rail. The subject then initiated the trial for the new rail when ready by pressing the joystick trigger switch. These procedures continued through the ninth rail. Subjects were free to rest for a while between rails, but seldom did. The same procedures and unblanking values for rails were used for perceptions of lateral distance and height, except that height trials started with unblanking at 1 m above the ground. For height left joystick up-down inputs were used for viewpoint motions and its trigger switch for starting the trial, and the right joystick top push button used to enter estimates. For lateral distance perceptions, subjects were required to set their eyepoint to the target value distance east of the east walls of the towers. For height perceptions, they set their eyepoint to the target value height above the ground.

Procedures for speed perceptions. The first trial for speed perception started on the highest rail with the helicopter

stationary and facing south at Y = 0. The #1 target speed for the increasing values sequence appeared on the panel CRT as the helmet display unblanked when the subject started the trial with the right joystick trigger switch. That joystick was used to accelerate to the speed perceived as equal to the first target speed, and the estimate entered with the left joystick top push button. The next target value appeared and estimation continued through the sixth increasing target value. When this estimate was entered the image was blanked to gray for 3 s during which the eyepoint was returned to the Y = 0 starting point, but at a speed upon unblanking three times the #6 target value, which also appeared again on the CRT as the initial value of the decreasing sequence. The subject reduced or increased speed to that perceived equal to the target value, and entered the estimate. Estimation continued through the last #1 target value estimate and image blanking. The eyepoint then was placed on the next lower rail under the same starting conditions, and the trial started by the subject when ready. Estimation continued through the fourth and lowest rail.

Subjects were allowed to take a break from testing at any time either in or out of the crew station. At the end of each set of rails, subjects were questioned as to their comfort and whether they desired a break. Tester set-up of new test conditions after a set of 9 or 4 rails always imposed at least a one minute break, during which subjects were encouraged to close their eyes and relax.

Results

With the magnitude production psychophysical procedure used, subject perceptions consist only of the six relative or six absolute target values for each of the four types of perception. Errors in perception consisted of the difference between a target distance or speed value to be perceived, and the magnitude of the viewpoint location or speed produced as the estimate for that value. Median accuracy of perception in percent of magnitude production was planned as the central tendency measure for analysis and presentation of results. However, some forward (Y) distance estimates were more than 100% under the target values. When used as denominators for converting magnitude production errors to perceptual accuracy errors, such values resulted in a non-continuous data distribution unsuitable for statistical analysis or meaningful interpretation. (They resulted in transformed errors that jumped from near plus to near minus infinity as magnitude production errors changed from just above to just below -100%.)

Consequently, magnitude production error in percent was used as the dependent variable measure for all data analyses. It was obtained by subtracting target value from magnitude production value, dividing this difference by the target value, and

multiplying by 100. This measure was used as the dependent variable for a set of nonparametric, correlation, multiple regression, and repeated measures ANOVA analyses of independent variable effects. Missing values of this measure were successfully replaced by BMDP program AM (Dixon, Brown, Engelman, & Jennrich, 1990). One or two extreme outliers for each type of perception also were replaced with the average of the two most extreme (for the first), or next two (for the second), non-outlier values.

Table 3 contains magnitude production percent error distribution statistics for the four types of perceptions. The min-max range for forward distance is about 70% greater than for lateral distance, and about three times greater than the ranges for height and speed. All the distributions are positively skewed and leptokurtic.

The distribution for speed perceptions is remarkable in that the minimum magnitude production error is 39% more than the target value. No magnitude productions less than the target value, as found for the other perceptions, are even approached for speed. Translating the 39% value into perceptual terms, the maximum percent perceived speed is only 72% of the simulated physical speed. All other speed perceptions are less than 72% of the physical speed.

Table 3

Magnitude Production Percent Error Distribution Statistics for Forward and Lateral Distance, Height, and Speed Perceptions

	Perception			
Statistic	Forward	Lateral	Height	Speed
Range	1714	1046	502	597
Minimum	-231	-82	-75	39
10th percentile	17	4	-11	78
25th percentile	69	44	7	105
Median	141	100	40	144
Mean	179	123	56	156
75th percentile	246	183	91	190
90th percentile	376	271	145	244
Maximum	1484	963	427	636
Skewness	1.907	1.172	1.230	1.678
Kurtosis	6.432	2.451	2.198	5.423

All observations involving each level of a factor were used in computing nonparametric and correlation statistics on the main effects of the factor, and multiple regressions for all factors

combined. Nonparametric tests included the sign, Wilcoxon signed rank, Friedman, and median tests (Norusis, 1993a). Repeated measures multi-factor ANOVAs (Norusis, 1993b) also were calculated to compare the other statistics with analyses of means-based main effects differences, and to assess interaction effects. For ANOVAs the log transform of magnitude production percent error was used, with 300 added to Y axis data, and 100 to X and Z data, to make all points positive prior to transformation. The log transformation corrected most of the positive skew and kurtosis in the original data distributions. Log transformed data also were used for correlations and multiple regressions. They increased most of the larger correlations over those for the original error data, and increased multiple R's by .02 to .06.

Statistical significance of factor main effects were closely comparable for all the nonparametric tests and correlations. Some of the ANOVA results were considerably less significant than the nonparametric and correlation statistics. The differences appeared to be due mainly to skew aspects which reduced mean log transformed level differences close to zero where larger median, rank and correlation differences existed. Probabilities from the median test were selected for use as the primary statistic for interpreting results because medians or geometric means usually are used in psychophysical research. While some reservations exist as to appropriateness of the median test for this repeated measures design, it is most directly related to the selected median measure of central tendency. It also gave slightly less significant probabilities than the other more powerful nonparametric and correlation tests.

Power Function Exponents

Psychophysical power function exponents and intercepts (Stevens, 1957) were computed using median magnitude productions as the physical parameters and target values as the perceptual parameters.

The power function exponent for forward distance perception is 1.00, for lateral distance 0.91, for height 0.88, and for speed perception 1.03. Respective zero intercepts are -0.76 m, -0.37 m, 0.41 m, and -4.56 km/hr. The 1.0 exponent for forward distance indicates overall changes in perceptions are exactly linearly proportional to median changes in physical magnitude productions. The 0.91 and 0.88 exponents for lateral distance and height perceptions indicate their proportions decrease as physical magnitudes increase. One should note the 1.0 power function exponent does not reflect accurate perception, but only changes linearly proportional to physical stimulus changes. For example, forward distance perceptions are found (see Figure 2) to be just 41% of the simulated physical distances produced for them.

Main Effects

Figure 2 shows median perceptions in percent of magnitude productions, for the four types of perceptions and for factor levels. It also shows the range of medians for individual pilots. The median percent perceptual accuracy values shown are direct transforms of the median percent error in magnitude productions that were used for statistical analyses. Accuracy values are used to facilitate subsequent comparison with real-world psychophysical data. Median percent perceptual accuracy for the levels of each factor are plotted as vertical lines connected by horizontal lines to form a bar. Length of the bar reflects the relative effect of each factor. Table 4 identifies the factor levels for these medians, and the median test probability for the main effect of factor level differences. Most median test probabilities for factor level differences are less than .00005.

Overall median perceptions are 41% of the simulated physical stimuli magnitude productions for forward distance and speed, 50% for lateral distance, and 72% for height perceptions.

For all four types of perceptions, helicopter flying experience significantly ($p < .00005$) improves perceptual accuracy over just limited fixed-wing flying experience. Helicopter experienced pilots are more accurate by 6% in forward distance perceptions, by 12% for lateral distance, by 13% for height, and by 7% for speed.

Accuracy for replications are about the same for forward and lateral distance perceptions, and, contrary to expectations, decreased on the second replication by 3% ($p = .0589$) for height and by 5% ($p < .00005$) for speed. The reduced accuracies on the second replication for height and speed are observed in the pilots with helicopter flying experience. Non-helicopter pilots actually improved slightly.

The best 3-D with familiar objects visual database level results in only a 4% ($p < .00005$) improvement in accuracy over just texture for forward distance perception, and only a 3% ($p = .0007$) improvement for lateral distance. For height and speed no significant differences between visual databases are found.

Relative perceptions are significantly ($p < .00005$) more accurate than absolute perceptions. The difference is 14% for forward distance, 15% for lateral distance, and 24% for height.

Less row offset of viewpoint rails results in statistically significant ($p < .00005$) improvements in accuracy for all four types of perceptions. The improvement in accuracy from the most to the least offset is 10% for forward distance, 13% for lateral

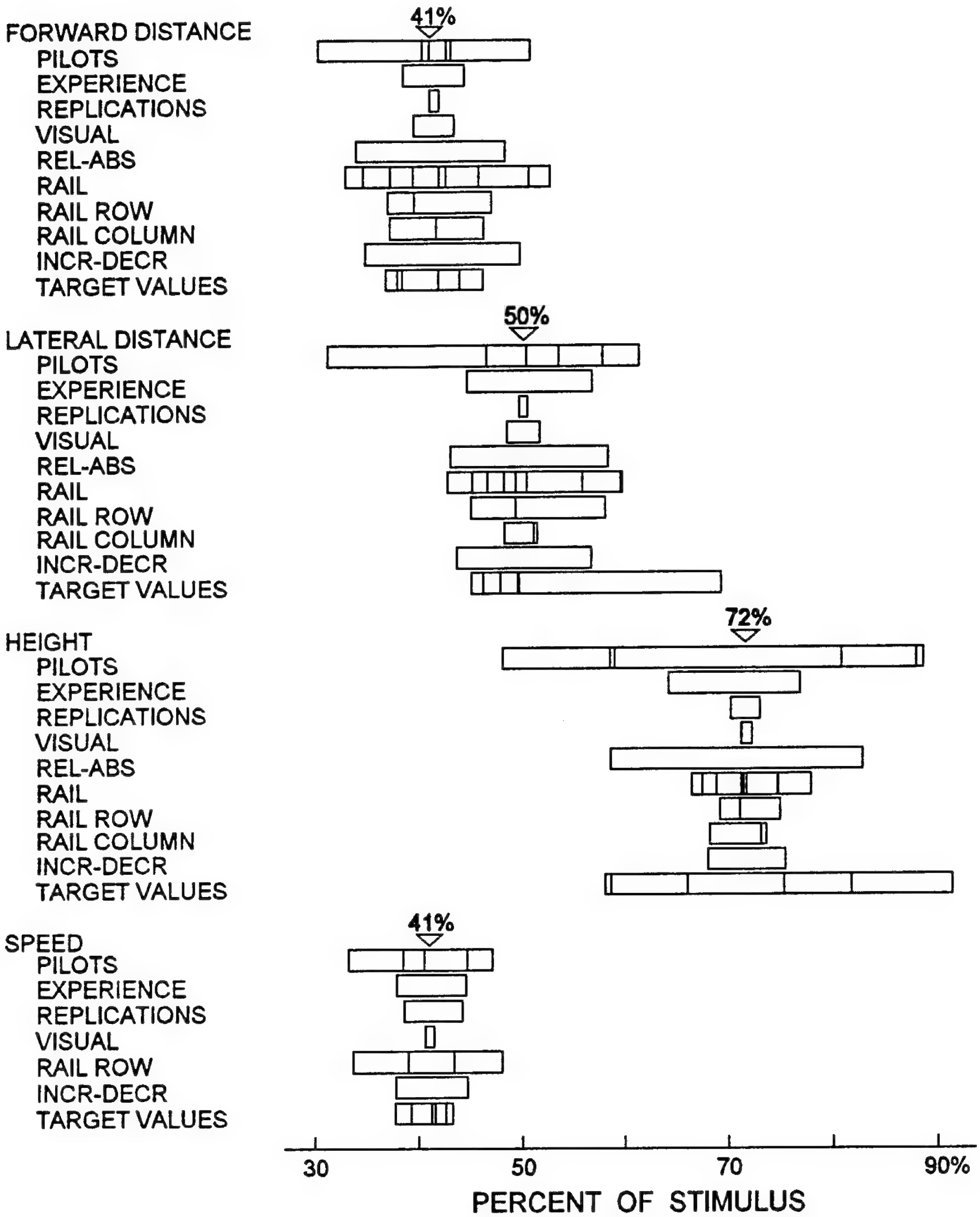

Figure 2. Median perceptions in percent of stimulus for type perception and factor levels. Bar vertical lines are factor levels.

Table 4

Median Perceptions in Percent of Physical Magnitude Productions, and Median Test Probabilities for Factor Level Differences

	Type of Perception			
Factors Levels	Forward Distance	Lateral Distance	Height	Forward Speed
Overall	41	50	72	41
Helicopter Flying	.0000[a]	.0000	.0000	.0000
Experience	44	57	77	45
None	38	45	64	38
Replication	.4874	.3593	.0589	.0000
First	41	50	73	44
Second	42	50	70	39
Visual	.0000	.0007	.3309	.2386
Best 3-D	43	52	72	41
Texture	39	49	71	42
Basis	.0000	.0000	.0000	
Relative	48	58	83	
Absolute	34	43	59	
Rail Row Offset	.0000	.0000	.0000	.0000
6 m	47	58	75	48
24 m	39	49	71	43
98 m	37	45	69	(24 m) 39
				(49 m) 34
Rail Column Offset	.0000	.0010	.0003	
6 m	46	51	74	
24 m	42	51	73	
98 m (forward 195)	37	48	68	
Sign of Change	.0000	.0000	.0000	.0000
Increase	50	57	75	45
Decrease	35	44	68	38
Target Values[b]	.0000	.0000	.0000	.0058
11-8-24	46	69	91	38
22-15-48	37	50	75	41
43-30-72	38	50	82	43
86-61-97	38	45	66	43
172-152-121	42	48	58	42
344-305-145	44	46	59	39

[a] Median test probability
[b] 1st relative and 2nd absolute in meters; 3rd speed in km/hr.

distance, and 14% for speed. These offsets all are in the vertical height dimension. For height, the least row offset results in 6% more accurate perceptions than the most, but this offset is in the X or lateral distance dimension.

Accuracy of forward distance perception improves by 9% from the most to the least rail column offset, but only by 3% for lateral distance and by 6% for height. All these differences are significant ($p < .00005$) statistically. For forward distance, the most offset is 195 m, but is 98 m for lateral distance and height. For height the column offset is in the Y dimension.

Increasing sequences of target values result in significantly ($p < .00005$) more accurate perceptions than decreasing sequences of target values. The improvement in accuracy is 15% for forward distance, 13% for lateral distance, and 7% for height and speed.

Significant differences in perceptual accuracy between target values exist for all four types of perceptions ($p < .00005$ for distances and height; $p = .0058$ for speed). The range between the least and most accurately perceived target values is 8% for forward distance, 24% for lateral distance, 33% for height, and 5% for speed.

The accuracy functions across the target values are different for each type of perception. For forward distance the function is concave upward, and it is concave downward for speed. For lateral distance it is relatively flat with an upward jump at the smallest target value. For height there is a roughly linear function of improved accuracy as target values decrease, except for a reversal at the next to smallest value. This appears to be due to alignment of the tops of the triangular relative reference cues on the towers for the #3 target value (1.0), which improves the accuracy of relative perceptions for it considerably.

Rails considered as a nine-level factor had a larger range of differences than when grouped by threes in either row or column offsets. The most accurate perceptions generally result from the least offset in both row and column, and the least accurate from the most offset in both. The range of difference in perceptual accuracy between least and most accurate rails ($p < .00005$) is 20% for forward distance, 17% for lateral distance, and 11% for height.

ANOVA probabilities on log mean differences between factor levels often are equal to the other statistical tests, but

this include the pilot experience factor for all four perceptions (p = .25 to .84), the visual factor for forward (.19) and lateral (.39) distance, the basis factor for lateral distance (.08), rail column offset for lateral distance (.14), and the sign of change factor for lateral distance (.04) and height (.10). The reason for these exceptions primarily appears to be due to skew-related reduction in factor level mean differences to values less than median differences. However, the greatly reduced degrees of freedom and the large mean square of the F test denominator for the between subjects experience factor probably is the reason for those exceptions.

Correlations and Multiple Regressions

Table 5 shows the independent variable multiple regressions and correlations with the logarithm of the dependent variable, percent error in magnitude productions. The SPSS multiple linear regression analysis program with stepwise variable selection was used for computing the multiple regressions and correlations (Norusis, 1993a). Independent variable factor levels were coded with sequential integer values for correlation and regression. Correlations and regressions were examined for rail offsets and target values coded with actual test values, found to be lower, and therefore were not used.

The maximum correlation for distances and height are with the basis (BAS) relative-absolute factor. For forward distance the BAS correlation is .346, for lateral distance .255, and for height .345. Due to the large N's, all correlations above .04 are statistically significant with probability less than .0005. For distances the second largest correlation is with the sign (SGN), target values increasing-decreasing factor. For height error the second largest correlation is with value (VAL) target values, and the third largest is with SGN. No other height correlations approach .1. For lateral distance perceptions experience (EXP), ROW, and VAL all are about .2. ROW is .184 and column (COL) .150 for forward distance. Speed error correlations reflect a different pattern, with ROW at -.508 the maximum. Other substantial speed correlations include EXP at .275, replication (REP) at .247, and SGN at .244. Only absolute speed perceptions were required. Therefore, the BAS factor is not applicable.

Multiple regression R's (see Appendix A) are .552 for forward distance, .499 for lateral distance, .476 for height, and .681 for speed. Variance (adjusted R square) accounted for by the regression equations are respectively, .304, .248, .225, and .461. These R square values indicate the independent variables

account for about one-fourth to one-half of magnitude production error variance.

Table 5

Factor Correlations (r) for Log of Percent Magnitude Production Error

Perception	EXP	REP	VIS	BAS	ROW	COL	SGN	VAL
	Independent Factors							
	Correlations[a]							
Forward Distance								
r	.043[b]	-.046	-.048	.346	.184	.150	.342	-.077
Lateral Distance								
r	.201	-.004[c]	-.051	.255	.199	.080	.232	-.205
Height								
r	.054	.040[d]	-.037[e]	.345	.057	.078	.121	.279
Speed								
r	.275	.247	-.002[f]	NA	-.508	NA	.244	-.098

Note. N for correlations is 5184, except for speed N = 1152. Full labels of the three letter abbreviations for the independent factors may be found in Table 1.
[a]One-tailed $p < .0005$ for all r's without superscript.
[b]$p = .001$. [c]$p = .397$. [d]$p = .002$. [e]$p = .004$. [f]$p = .479$.

While the stepwise increases in multiple R generally are proportional to factor correlation values, some of the factor weights are not at all proportional. Weights for forward distance are closely proportional. The VAL weight for lateral distance is not at all proportional. SGN and VAL weights for height are not proportional, although they retain an ordinal relationship. For speed, the ROW correlation is about twice as large as the other three substantial correlations (EXP, REP, and SGN), yet its weight is less than theirs.

Interactions

Interactions between the factors in this experiment far exceed that expected on a chance basis, and are too numerous to

cover in detail. Excluding interactions with the pilot experience factor, there are 121 interactions in each of the ANOVA analyses for distances and height. By chance, 6 interactions of less than .05 probability could be expected, 1.2 interactions of less than .01, and .12 less than .001. For forward distance, 26 interactions less than .05 were obtained, 18 less than .01, and 10 less than .001. For lateral distance, the respective numbers were 22, 15 and 5, and for height 22, 5 and 4. The ratios of obtained to expected significant interactions were greater for 2-way and 3-way interactions than for 4-way and higher interactions. In addition, there were about the same number of significant interactions which involved the pilot experience factor, although most of them were in the .05 to .01 range.

These much larger than expected numbers of significant interactions suggest that adoption of certain combinations of factor levels can be expected to result in much better perceptual accuracy than for other combinations. The effects of such combinations were explored with the combination of the best rail with the best level of the non-rail factors over all target values. The worst such combination also was examined. For the best levels median forward distance perceptions were found to be 96% of physical magnitude productions, lateral distance and height perceptions 99%, and speed perceptions 60%. These all represent substantial improvements over the median accuracy for all levels of all factors combined (41, 50, 72 and 41%, respectively), and three of them are close to perfectly accurate. The worst combination was found to reduce respective accuracies to 25, 39, 51 and 32%.

The best combination of factor levels, however, used the increasing sequence of target values, which involves moving backward for forward distance perceptions, away from the reference for lateral distance, and upward for height. In flying and most other activities, critical forward distance perceptions must be made while moving forward, and lateral distance and height perceptions made while moving toward an object or the ground. The decreasing level of change in target values was substituted, therefore, to reflect these important aspects of applied aviation distance and height perception. With this substitution in the best combination of other factor levels, median forward distance accuracy dropped to 48%, lateral distance to 67%, height accuracy remained at 99%, and speed accuracy dropped to 55%.

Discussion

Simulator Versus Real-World Perceptions

The primary issue of concern in simulator or virtual reality psychophysics is how well perceptions with the visual system match those in the real-world. These results were not encouraging. Only height perceptions appear to be reasonably comparable with the accuracy of perceptions in the real-world. Real world relative perceptions have been found to be more accurate than absolute perceptions (Gibson, Bergman & Purdy, 1955), and this also was found for these simulator distance and height perceptions. Relative height perceptions at 83% of actual stimuli, in particular, were reasonably accurate.

Median absolute forward distance perceptions with the simulator visual system were 34% of simulated physical distance, while typical real-world absolute perceptions were about 91% of actual physical distances (Gibson & Bergman, 1954). Lateral absolute distance perceptions in the simulator of 43% of physical distances also were much less than for real-world perceptions.

Median absolute height perceptions with simulator vision were 59% of simulated physical heights, while real-world absolute perceptions over land average 87% of actual heights (Ungs & Sangal, 1990).

Median speed perceptions with simulator vision were 41% of simulated physical speeds, while typical real-world perceptions of speed are 90% of actual (Salvatore, 1968).

Reasons for the large differences in accuracy between simulator and real world perceptions of distance, height and speed are not evident. It was hoped the high resolution, wide field of view, head-slaved, color virtual reality type of display used in this research would result in close to veridical perceptions. This did not occur. Characteristics of the display image, image content, selected test factors and levels, test procedures, or interactions between these aspects, must be causing the differences. The current results suggest several causal factors that may explain some part of the differences, but are far from providing a complete explanation.

There are trends of more accurate simulator perceptions as eye height is reduced. The minimum eye height used of 6 m is about four times the 1.5 m eye heights used in the real world distance and the primary speed perception tasks used for comparison. The eye height trends suggest perceptions at 1.5 m eye height ought to be about 10% more accurate for forward and

lateral distance and speed, perhaps more so for the distances. The larger viewpoint rail column offsets also had a negative impact on perceptual accuracies. Being closer to, or moving more directly toward or from the perceptual reference, resulted in slightly more accurate perceptions.

The minimal improvement over texture in accuracy of distance perceptions with the high detail 3-D visual database with familiar objects, and the lack of improvement in height and speed perceptions, was not expected. Lack of counterbalancing of order of testing texture and 3-D visuals may have obscured an actual effect. The 3-D database, with the same texture of the texture database, was always tested before the texture only condition. This may have resulted in perceptual calibration of the texture under the 3-D condition, and masked actual differences.

Rationale for the 3-D condition first was that opportunity for exposure in flight training to a detailed 3-D part of the database, will almost always exist before having to fly over the texture only database areas. The 3-D condition first addresses the applied issue of acceptability of texture for flight training, but may have obscured the more fundamental issues of the effects of visual database detail and content familiarity. This minimal impact of visual scene detail may be of substantial consequence from a practical standpoint in visual system design and application. One should note, however, that Kleiss and Hubbard (1993) found increasing density of vertical objects improved altitude change perceptions near the ground surface independently of ground texture. It should not be surprising that different visual scene characteristics may be required to support the different perceptual discriminations involved in flying.

Conclusions

It was hoped that very high quality pilot's virtual reality type of visual display which was used might reduce or eliminate the large perceptual errors commonly observed anecdotally for most flight simulator visual display systems. This hope was not realized. While similar psychophysical data for other simulator visual display systems do not exist to provide definitive comparisons, these results suggest there may not be large differences. If so, they suggest that head slaving a high quality very wide field of view display has contributed little to improving the match of simulator visual perceptions with those in the real world, except perhaps for height perceptions. The failure to produce a close match with real world perceptions is

logically due to some combination of image display characteristics and image content, or to differences in test conditions.

From a methodological standpoint the results of this psychophysical research are encouraging. Large and statistically significant differences were found between most of the experimental design factors, and between many of their interactions. Although practical significance may be debatable, the large number of observations resulted in high statistical significance for correlations of .04 or greater, and differences in perceptual accuracy of 3% or greater. The magnitude production psychophysical method used appears to be sensitive to the characteristics of visual scene content and user tasks. These have proven difficult to assess precisely by means of more global measures of flying or combat performance.

References

Armstrong, R. N., Hofmann, M. A., Sanders, M. G., Stone, L. W., & Bowen, C. A. (1975). Perceived velocity and altitude judgments during rotary wing aircraft flight (Report 76-3). Fort Rucker, AL: U.S. Army Aeromedical Research Laboratory.

Barrette, R., Dunkley, K., Kruk, R., Kurts, D., Marshall, S., Williams, T., Weissman, P., & Antos, S. (1990). Flight simulator: Advanced wide field-of-view, helmet-mounted, infinity display system. (AFHRL-TR-89-36). Williams Air Force Base, AZ: CAE Electronics Ltd., Martin Shenker Optical Design Inc., and Schott Fiber Optics, Inc.; for Operations Training Division, Air Force Human Resources Laboratory.

Buckland, G. H., Monroe, E. G., & Mehrer, K. I. (1980). Flight simulator visual textural cues for landing (AFHRL-TR-79-81). Williams Air Force Base, AZ: Operations Training Division, Air Force Human Resources Laboratory.

Buckland, G. H., Edwards, B. J., & Stephens, S. (1981). Flight simulator visual development and instructional features for terrain flight simulation. Proceedings of the 1981 Image Generation/Display Conference II (pp. 351-362). (AFHRL-TR-81-48). Williams Air Force Base, AZ: Operations Training Division, Air Force Human Resources Laboratory.

Burney, G. M. (1977). Estimation of distances while driving (Supplementary Report 262). Crowthorne, Berkshire, England: Department of Transport, Department of the Environment, Transport and Road Research Laboratory.

De Maio, J., & Brooks, R. (1982). Assessment of simulator visual cuing effectiveness by psychophysical techniques. Proceedings of 4th Interservice/Industry Training Equipment Conference (pp. 379-381). Orlando, FL.

De Maio, J., Rinalducci, E. J., Brooks, R., & Brunderman, J. (1983). Visual cueing effectiveness: Comparison of perception and flying performance. Proceedings of 5th Interservice/ Industry Training Equipment Conference (pp. 92-96). Washington, DC.

Denz, E. A., Palmer, E. A., & Ellis, S. R. (1980). Effect of field of view and monocular viewing on angular size judgments in an outdoor scene. (TM 81176). Moffett Field, CA: Ames Research Center, NASA.

Dixon, W. J., Brown, M. B., Engelman, L., & Jennrich, R. I. (Eds.). (1990). BMDP statistical software manual. Los Angeles: University of California Press.

Fine, B. J., & Kobrick, J. L. (1981). Individual differences in distance estimation: Comparison of judgments in the field with those from projected slides of the same scene. (Report M 7/82). Natick, MA: U.S. Army Research Institute of Environmental Medicine.

Galanter, E., & Galanter, P. (1973). Range estimates of distant visual stimuli. Perception & Psychophysics, 14(2), 301-306.

Gibson, E. J., & Bergman, R. (1954). The effect of training on absolute estimation of distance over the ground. Journal of Experimental Psychology, 48, 473-482.

Gibson E. J., Bergman, R., & Purdy, J. (1955). The effect of prior training with a scale of distance on absolute and relative judgments of distance over ground. Journal of Experimental Psychology, 50, 997-1005.

Gilinsky, A. S. (1951). Perceived size and distance in visual space. Psychological Review, 58, 460-482.

Gilinsky, A. S. (1955). The effect of attitude upon the perception of size. American Journal of Psychology, 68, 173-192.

Kennedy, R. S., Berbaum, K. S., Collyer, S. C., May, J. G., & Dunlap, W. P. (1988). Spatial requirements for visual simulation of aircraft at real-world distances. Human Factors, 30(2), 153-161.

Kleiss, J. A., & Hubbard, D. C. (1993). Effects of three types of flight simulator visual scene detail on detection of altitude change. Human Factors, 35(4), 653-671.

Kling, J. W., & Riggs, L. A. (1971). Experimental psychology. New York: Holt, Rinhart, and Winston.

Kraft, C. L., Anderson, C. D., & Elworth, C. L. (1980). Psychophysical criteria for visual simulation systems (AFHRL-TR-79-30). Williams Air Force Base, AZ: Boeing Aerospace Company, for Operations Training Division, Air Force Human Resources Laboratory.

Kraft, C. L., & Anderson, C. D. (1980). Psychophysical criteria for visual simulation systems: Phase II—Experimental investigations of display joints and scene inserts (AFHRL-TR-80-18). Williams Air Force Base, AZ: Boeing Aerospace Company, for Operations Training Division, Air Force Human Resources Laboratory.

Lintern, G., Thomley-Yates, K. E., Nelson, B. E., & Roscoe, S. N. (1987). Content, variety and augmentation of simulated visual scenes for teaching air-to-ground attack. Human Factors, 29(1), 45-59.

McCormick, D., Smith, T., Lewandowski, F., Preskar, W., & Martin, E. (1983). Low-altitude database development evaluation and research (LADDER). Proceedings, 5th Interservice/Industry Training Equipment Conference (Vol I, pp. 150-155). Washington, DC.

Martin, E. L., & Rinalducci, E. J. (1983). Low-level flight simulation: Vertical cues (AFHRL-TR-83-17). Williams Air Force Base, AZ: Operations Training Division, Air Force Human Resources Laboratory.

Norusis, M. J. (1993a). SPSS for Windows: Base system users guide, Release 6.0. SPSS, Inc., Chicago, IL.

Norusis, M. J. (1993b). SPSS for Windows: Advanced statistics, Release 6.0. SPSS, Inc., Chicago, IL.

Palmer, E., & Petitt, J. (1977). A measure of psychological realism on a visual simulator. Journal of Aircraft, 14(5), 421-422.

Salvatore, S. (1968). The estimation of vehicular velocity as a function of visual stimulation. Human Factors, 10(1), 27-32.

Stevens, S. S. (1957). On the psychophysical law. Psychological Review, 64, 153-181.

Teghtsoonian, R., & Teghtsoonian, M. (1970). Scaling apparent distance in a natural outdoor setting. Psychonomic Science, 21(4), 215-216.

Ungs, T. J., & Sangal, S. P. (1990). Perception of near earth altitudes by pilots: Ascending vs. descending over both a land and water surface. Aviation, Space and Environmental Medicine, 61, 1098-1102.

APPENDIX A

Multiple Regressions and Factor Weights

APPENDIX A

Table A-1

Multiple Regressions and Factor Weights for Log of Percent Magnitude Production Error

	Independent Factors							
	EXP	REP	VIS	BAS	ROW	COL	SGN	VAL
Forward Distance								
Multiple R: .552								
Constant: 2.545								
Weight	.011	-.012	-.013	.093	.030	.025	.091	-.006
Step Entered	8	7	6	1	3	4	2	5
R After Entry	.552	.551	.549	.346	.520	.541	.487	.547
Lateral Distance								
Multiple R: .499								
Constant = 2.106								
Weight	.087		-.022	.111	.053	.021	.101	.026
Step Entered	4	Not	7	1	5	6	2	3
R After Entry	.448		.499	.255	.490	.497	.345	.401
Height								
Multiple R = .476								
Constant = 1.972								
Weight	.020	.015	-.013	.127	.013	.017	.044	.030
Step Entered	6	7	8	1	5	4	3	2
R After Entry	.473	.474	.476	.345	.470	.466	.460	.443
Speed								
Multiple R: .681								
Constant: 2.268								
Weight	.107	.096		NA	-.088	NA	.095	-.011
Step Entered	2	3	Not	NA	1	NA	4	5
R After Entry	.577	.628		NA	.508	NA	.674	.681

Note. N for multiple regressions is 5184, except 1152 for speed. Full labels of the three-letter abbreviations for the independent factors may be found in Table 1.